

了解ACI交换矩阵中的PXE引导并排除故障

目录

[预引导执行环境 \(PXE, 有时读作pixie\)](#)

[将PXE客户端重定向到不同的TFTP服务器](#)

[重要故障排除说明](#)

[PXE引导中可能使用的DHCP选项](#)

[客户拓扑](#)

预引导执行环境(PXE, 有时发音为pixie)

PXE或网络引导允许BIOS通过网络下载和引导操作系统(OS)。在Cisco NX-OS路由器中, 这相当于引导加载程序[loader>]ak Kickstart。

在PXE引导中, PXE客户端的BIOS会查看配置的引导顺序, 如果配置了PXE引导, 则确定硬件是否支持PXE引导。如果是, 它会尝试将IP地址从NIC上DHCP。在DHCP请求中, PXE客户端请求某些DHCP选项, 如67 [文件名]。在DHCP应答中, DHCP服务器使用请求的选项进行应答。然后, PXE客户端下载在DHCP响应中指定的文件名[网络引导程序(NBP)], 并运行它以获取加载到服务器上的操作系统。

将PXE客户端重定向到不同的TFTP服务器

默认情况下, PXE客户端将尝试从DHCP服务器下载DHCP响应中的文件名, 除非它在DHCP响应中收到选项66、下一服务器或选项60/43以将其重定向到其他IP地址。下一个服务器地址用于bootp和PXE, 以指定将不同的服务器用于DHCP和TFTP。

next-server (选项12) 和bootfile (选项0) DHCP选项通常指定TFTP服务器和初始引导加载程序的文件名。Bootfile (可选) 由客户端在DHCPDISCOVER消息中请求特定类型的引导文件。DHCPPOFFER中的DHCP服务器使用它来完全指定引导文件目录路径和文件名。

还必须为文件名指定选项67。bootfile和选项67之间的区别是在DHCP响应中找到文件名的位置。使用保留字段, 响应不会在响应数据包的“选项”部分进行编号, 而是按特定顺序给出。

例如, 服务器主机名和引导文件显示在下图中的编号选项之前。

```

▶ Internet Protocol Version 4, Src: 10.1.120.91 (10.1.120.91), Dst: 10.1.208.93 (10.1.208.93)
▶ User Datagram Protocol, Src Port: 67 (67), Dst Port: 67 (67)
▼ Bootstrap Protocol (Offer)
  Message type: Boot Reply (2)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 0
  Transaction ID: 0xb6c37a01
  Seconds elapsed: 0
  ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 10.250.118.251 (10.250.118.251)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 10.250.118.1 (10.250.118.1)
  Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name: /pxelinux.0
  Magic cookie: DHCP
  ▼ Option: (53) DHCP Message Type (Offer)
    Length: 1
    DHCP: Offer (2)
  ▼ Option: (54) DHCP Server Identifier
    Length: 4
    DHCP Server Identifier: 10.250.120.2 (10.250.120.2)
  ▼ Option: (51) IP Address Lease Time
    Length: 4
    IP Address Lease Time: (85537s) 23 hours, 45 minutes, 37 seconds
  ▼ Option: (58) Renewal Time Value
    Length: 4
    Renewal Time Value: (42768s) 11 hours, 52 minutes, 48 seconds

```

重要故障排除说明

如果PXE引导客户端未在DHCP请求数据包的选项55参数请求列表中请求特定DHCP选项，则DHCP选项将不会在DHCP服务器响应数据包中，即使它在DHCP服务器上配置。获取嗅探器跟踪以验证请求和提供的选项非常重要。

NOTE:如果为DHCP服务器的DHCP ip helper地址 (DHCP中继) 配置了ACI，则第一跳枝叶交换机上可能会使用tcpdump。由于枝叶交换机必须充当DHCP中继，因此CPU必须处理DHCP数据包，将其转发到DHCP服务器。您可以使用管理员帐户登录，然后使用命令

```
tcpdump -i any -w /tmp/packet.pcap "port 67"
```

捕获主机和DHCP服务器之间的DHCP数据包。如果枝叶交换机没有任何OOB ip地址，您可以使用管理员用户名登录APIC并使用

```
scp admin@<switch TEP>:/tmp/packet.pcap /home/admin
```

然后，您可以使用GUI IP地址从APIC传输文件。

根据PXE服务器支持的文件名，可能有许多不同的文件名。Linux服务器最常见的引导加载程序文件名称为“/pxelinux.0”，通常位于/tftpboot目录中。

PXE引导中可能使用的DHCP选项

名称	选项编号	配置
引导文件	0	Bootfile <file>
DEFAULT_ROUTER	3	Default-router 10.250.118.1 Network
网络	11	10.250.118.0 255.255.255.0
NEXT_SERVER	12	下一台服务器 10.45.67.8
使用匹配选项60发送	43	选项43 hex 06010708070000010afbd015
选项43的匹配字符串	60	选项60 "PXEClient"
单个TFTP服务器	66	选项66 ip 10.240.221.21
PXE文件名	67	选项67 "/pxelinux.0"
TFTP服务器列表	150	选项150 10.240.208.21

如果DHCP服务器将其选项60字符串与客户端发送的选项60字符串匹配，则选项43将返回给PXE客户端。选项43有几个子选项（选项6和8等）：

子选项6定义为

06：子选项6

01：1字节长

07：使用PXE_BOOT_SERVERS列表，禁用组播和广播发现（仅单播）

子选项8定义为

08：子选项8

07：08选项的大小（字节）

0000：PXE服务器类型0（PXE引导服务器）

01：PXE服务器数

0afbd015:10.251.208.21

该选项在DHCP配置中类似

选项43 hex 06010708070000010afbd015

根据<https://support.microsoft.com/en-us/kb/259670>，当DHCP服务器设置了以下选项时：

60 =客户端标识符（设置为“PXEClient”）

66 =引导服务器主机名或IP地址

67 =引导文件名

当来自DHCP服务器的初始DHCP提供包含这些引导选项时，会尝试从PXE客户端连接到DHCP服务器上的端口4011，并出现“代理DHCP服务未在端口4011上回复”错误。

NOTE:Microsoft不支持在DHCP服务器上使用这些选项重定向PXE客户端。

客户拓扑

在本示例中，用户有2个不同的ACI交换矩阵。在ACI交换矩阵#1中，PXE客户端正在DHCPing和引导交换矩阵内的Microsoft服务器。在另一个ACI交换矩阵中，PXE客户端通过L3Out从外部IOS-XE 4507路由器DHCP连接，但使用同一Microsoft服务器下载文件，以便DHCP服务器不充当TFTP服务器。

客户最终采用的解决方案是迁移到Linux DHCP服务器。分析DHCP数据包跟踪时，客户似乎误配置了选项43字符串，IOS-XE DHCP服务器从未工作。

要在ACI中运行，请执行以下操作：

1. 必须配置DHCP中继以中继DHCP客户端数据包
2. 合同必须允许L3Out和PXE客户端所在的EPG之间的DHCP和TFTP数据包

在枝叶交换机上，您可能会看到使用命令的TFTP数据包的数据包丢弃

```
calo2-leaf2# show logging ip access-list internal packet-log | grep <client ip address>
```

此命令将在2.0版中更改，以添加deny或permit关键字，并变为

```
calo2-leaf2# show logging ip access-list internal packet-log deny | grep <client ip address>
```

在示例中，PXE客户端未请求选项66，而是在选项5参数列表中指定选项60并请求选项43。它也在请求下一台服务器（选项12）。请注意，它通过选项67请求了PXE文件名。

```

▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
Client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 0.0.0.0 (0.0.0.0)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 10.250.118.1 (10.250.118.1)
Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
Client hardware address padding: 00000000000000000000
Server host name not given
Boot file name not given
Magic cookie: DHCP
▶ Option: (53) DHCP Message Type (Discover)
▼ Option: (55) Parameter Request List
  Length: 24
  Parameter Request List Item: (1) Subnet Mask
  Parameter Request List Item: (2) Time Offset
  Parameter Request List Item: (3) Router
  Parameter Request List Item: (5) Name Server
  Parameter Request List Item: (6) Domain Name Server
  Parameter Request List Item: (11) Resource Location Server
  Parameter Request List Item: (12) Host Name
  Parameter Request List Item: (13) Boot File Size
  Parameter Request List Item: (15) Domain Name
  Parameter Request List Item: (16) Swap Server
  Parameter Request List Item: (17) Root Path
  Parameter Request List Item: (18) Extensions Path
  Parameter Request List Item: (43) Vendor-Specific Information
  Parameter Request List Item: (54) DHCP Server Identifier
  Parameter Request List Item: (60) Vendor class identifier
  Parameter Request List Item: (67) Bootfile name
  Parameter Request List Item: (128) DOCSIS full security server IP [TODO]
  Parameter Request List Item: (129) PXE - undefined (vendor specific)
  Parameter Request List Item: (130) PXE - undefined (vendor specific)
  Parameter Request List Item: (131) PXE - undefined (vendor specific)
  Parameter Request List Item: (132) PXE - undefined (vendor specific)
  Parameter Request List Item: (133) PXE - undefined (vendor specific)
  Parameter Request List Item: (134) PXE - undefined (vendor specific)
  Parameter Request List Item: (135) PXE - undefined (vendor specific)

```

响应显示bootfile和选项67

- ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
Client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 10.250.118.251 (10.250.118.251)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 10.250.118.1 (10.250.118.1)
Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
Client hardware address padding: 0000000000000000000000
Server host name not given
Boot file name: /pxelinux.0
Magic cookie: DHCP
- ▶ Option: (53) DHCP Message Type (Offer)
- ▶ Option: (54) DHCP Server Identifier
- ▶ Option: (51) IP Address Lease Time
- ▶ Option: (58) Renewal Time Value
- ▶ Option: (59) Rebinding Time Value
- ▶ Option: (1) Subnet Mask
- ▶ Option: (6) Domain Name Server
- ▶ Option: (3) Router
- ▼ Option: (67) Bootfile name
 Length: 11
 Bootfile name: /pxelinux.0
- ▶ Option: (82) Agent Information Option
- ▶ Option: (255) End

枝叶交换机上的TCPdump: